iEC7871 *Quercus suber* model: the first multi-tissue diel cycle genome-scale metabolic model of a woody tree

Emanuel Cunha[a], Miguel Silva[a], Ines Chaves[b], Huseyin Demirci[a], Davide Lagoa[a], Diogo Lima[a], Miguel Rocha[a], Isabel Rocha[b,1,2], and Oscar Dias[a,1,2]

[a] Centre of Biological Engineering, Universidade do Minho, 4710-057 Braga, Portugal

[b] Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Avenida da República, Quinta do Marquês, 2780-157 Oeiras, Portugal

[1] Senior authors

[2] Corresponding authors: irocha@itqb.unl.pt; odias@ceb.uminho.pt

Abstract

In the last decade, genome-scale metabolic models have been increasingly used to study plant metabolic behaviour at the tissue and multi-tissue level in different environmental conditions. *Quercus suber* (*Q. suber*), also known as the cork oak tree, is one of the most important forest communities of the Mediterranean/Iberian region. In this work, we present the genome-scale metabolic model of the *Q. suber* (iEC7871), the first of a woody plant. The metabolic model comprises 7871 genes, 6230 reactions, and 6481 metabolites across eight compartments. Transcriptomics data was integrated into the model to obtain tissue-specific models for the leaf, inner bark, and phellogen. Each tissue's biomass composition was determined to improve model accuracy and merged into a diel multi-tissue metabolic model to predict interactions among the three tissues at the light and dark phases. The metabolic models were also used to analyze the pathways associated with the synthesis of suberin monomers. Nevertheless, the models developed in this work can provide insights about other aspects of the metabolism of *Q. suber*, such as its secondary metabolism and cork formation.

1. Introduction

The cork oak, *Quercus suber* L., is a characteristic tree of the Mediterranean/Iberian landscape ecosystem. The tree forms a thick bark of cork (phellem) containing high levels of aliphatic suberin (referred to as suberin), aromatic suberin (lignin), extractives (waxes and tannins), and polysaccharides (Pereira, 1988a). Cork properties are so unique that this polimer is used to deal with extreme conditions such as the ones faced by spacecrafts, providing insulation solutions to NASA and the European Space Agency (ESA). The aliphatic suberin component has particular interest since it provides waterproof, light, elastic, and fire-retardant properties to cork (Vaz et al., 2011; Pereira,

2015). Only the outher bark is separated from the trunk during the harvest, which enables the regeneration and allows using the tree as a renewable biological resource. The phellem is produced by the phellogen (cork cambium), involving the proliferation of phellogen derivates cells which undergo to differentiate into cork cells through cell expansion, deposition of suberin and waxes, and an irreversible senescence program ending with cell death. The first debarking occurs when the tree has 18-25 years old. This first harvested cork , known as "virgin" cork, produced from the original phellogen, yields poor-quality cork. The original phellogen is then replaced by a traumatic phellogen that proliferates to originate a new cork layer. From the third debarking onwards, a higher quality cork with high economic value is obtained - reproduction cork (or "amadia") (Graça and Pereira, 2004). Cork growth and quality is dependent on the genome of each tree but is also strongly dependent on the environment, such as water availability, temperature, pests and diseases.

The cork oak tree is cultivated in Portugal, Spain, Algeria, Morocco, France, Italy, and Tunisia, having a substantial economic impact on these countries. This tree species can generally live over 200 years, which implies a capacity to bear with biotic stresses like bacteria, fungi and insects, as well as abiotic stresses like drought, floods and fires (Acácio et al., 2007; Kim et al., 2017).

Due to its economic interest for Portugal, a national consortium was created to first sequence the transcriptome (Pereira-Leal et al., 2014) and recently the full sequencing of the genome of *Quercus suber* (Ramos et al., 2018). The genome sequence's availability allows the application of systems biology tools to study this species' metabolic behaviour.

Genome-scale metabolic (GSM) models aim at depicting the whole metabolic network of an organism. Such models have been widely used for metabolic engineering purposes, mainly with prokaryotes and yeasts. Among other applications, GSM models can be used to analyze an organism's metabolic behaviour in different environmental and genetic conditions, including the effect of genes knock-outs and over/under-expression (Dias and Rocha, 2015). Curated models have proven to accurately predict complex organisms' metabolic behaviour in diverse areas of knowledge, from biotechnological or environmental to medical applications (Fong et al., 2005; Bordbar et al., 2011; Agren et al., 2014; Jerby-Arnon et al., 2014; Zhang and Hua, 2016). Genome-scale modelling in plants is much challenging than in prokaryotes. The struggle starts in the annotation of the complex genomic content, in which the purpose of a significant number of genes is still unknown. Also, the subcellular compartmentalization, tissue differentiation, and interactions between tissues are complex in plants. Thus, the utilization of user-friendly tools designed for this purpose, like *metabolic models reconstruction using genome-scale information* (*merlin)* (Dias et al., 2015), accelerates the reconstruction process. In the last decade, the number of published GSM models of plants has

increased considerably. Organisms, like *Arabidopsis thaliana* (Dal'Molin et al., 2010; Cheung et al., 2013), *Zea mays* (maize) (Saha et al., 2011; Seaver et al., 2015), *Oryza sativa* (rice) (Lakshmanan et al., 2013; Poolman et al., 2013) have more than one model available. Considering the importance of plants in terms of nutrition, biofuels, and their capability to produce a variety of secondary metabolites, it is not surprising that studies for plant genome-scale model reconstruction will increase, parallel to the growth in the number of sequenced species.

This paper is the first reconstructing a GSM model for the cork oak tree, integrating tissue-specific models, and a multi-tissue metabolic model that can be used to study the metabolic behaviour of *Q. suber*. As far as we know, this is the first published GSM model of a tree.

1 Results

1.1 Genome Annotation

A GSM model for *Q. suber* based on an up-to-date genome annotation was reconstructed in this work. A Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) search against the Swiss-ProtKB (Apweiler et al., 2011) database allowed to identify similarity results for 47,199 out of the 59,614 genes available in the genome (Ramos et al., 2018). A second BLAST search against UniProtKB/TrEMBL allowed obtaining hits for 12,415 genes, while 590 had no results.

As detailed in the Materials and Methods section, *merlin*'s"automatic workflow" tool annotates genes based on the homologous gene records taxonomy. Most enzyme-encoding genes were annotated based on *A. thaliana* gene records (82%). Organisms not included in the annotation workflow represent 16% of gene annotations, and only 1% were annotated using *Quercus* genus' gene records. A more detailed analysis of the genome annotation is available in Supplementary File 1.

1.2 Model Properties

The metabolic model is mass balanced and can predict growth in phototrophic and heterotrophic conditions. These conditions were defined by setting the photon and CO_2 (or sucrose) as the sole energy and carbon source, respectively. Additionally, the model requires NH_3 or HNO_3, H_3PO_4 , H_2SO_4, Fe^{2+}, and Mg^{2+} to create biomass. The uptake of water and H^+ was also allowed. Photorespiration was tested using photon and CO_2 as energy and carbon source, respectively, and constraining the carboxylation/oxygenation ratio activity of Rubisco (reactions R00024 and R03140) to 3:1. The model was tested to guarantee that no biomass or energy is produced without energy input in each condition.

The general properties of the cork oak model and other five published plant models - arabidopsis (Chung et al., 2013; Seaver et al., 2015), maize (Saha et al., 2011; Seaver et al., 2015), and soybean (Moreira et al., 2019) - are presented in table 1.

Table 1: General properties of the Cork Oak model, and other four published plant GSM models. Only generic models were considered.

	Quercus suber (this work)	*Arabidopsis thaliana* (Chung et al., 2013)	*Arabidopsis thaliana* (Seaver et al., 2015)	*Zea mays* (Saha et al., 2011)	*Zea mays* (Seaver et al., 2015)	*Glycine max* (Moreira et al., 2019)
Genes	7871	1475	5184	1563	13,279	6127
Reactions	6230	1895	6399	1970	6458	2984
Metabolites	6481	1761	6236	2129	6250	2814
Compartments	8	6	9	6	9	5

The Cork Oak model comprises 7871 genes, 6230 reactions (enzymatic and transport), 708 exchange reactions, and 6476 metabolites distributed across eight subcellular compartments – the cytoplasm, mitochondria, plastid, endoplasmic reticulum, Golgi apparatus, vacuole, peroxisome, and the extracellular space. The number of genes, reactions, and metabolites of the Cork Oak model are only comparable with the *A. thaliana* and *Z. mays* models developed by Seaver (Seaver et al., 2015).

The reactions and metabolites present in the cork oak model were compared with the models developed by Chung for *A. thaliana* (Chung et al., 2013), and Saha for Z. mays (Saha et al., 2011) (Figure 1 and Supplementary File 2), as the reactions and metabolites available in these models have with the same identifiers (KEGG) as our model.

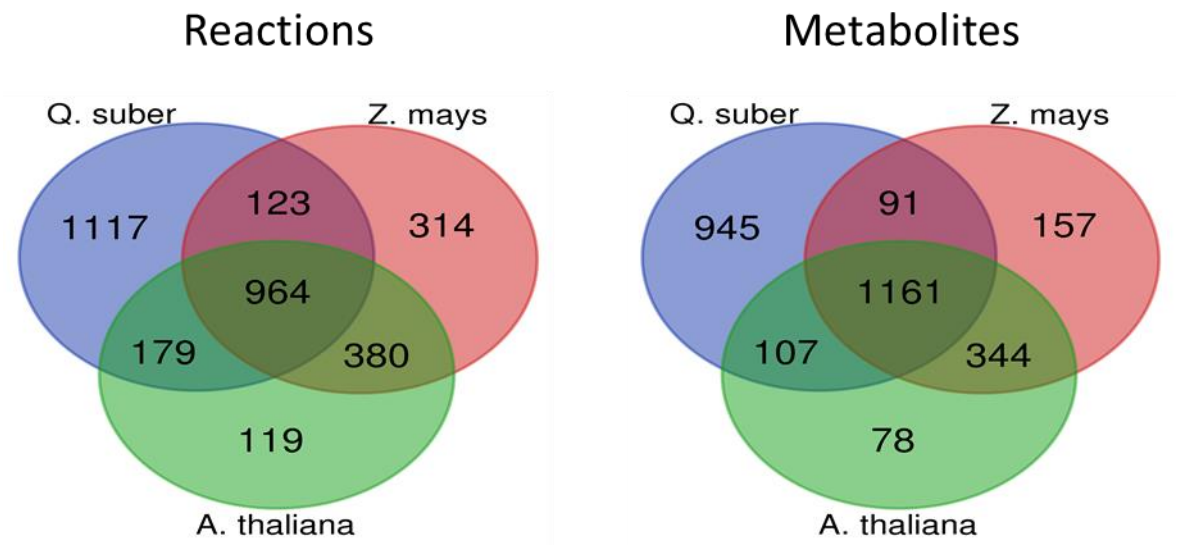

Figure 1: Venn diagram of reactions (left) and metabolites (right) included in the Q. suber, A. thaliana, and Z. mays models.

A total of 3196 unique reactions and 2883 unique metabolites were identified across the three models. The cork oak model includes 1117 reactions that are not present in the other two. It has 123 reactions in common only with the Maize model and 179 only with the *A. thaliana* model. The maize and *A. thaliana* models share 380 reactions, and 964 were found in the three models. The metabolites analysis shows similar behaviour. This analysis is detailed in Supplementary File 2 Tables S4-S5.

The reactions only available in the Cork Oak model were further analyzed to assess the annotation of the genes and pathways associated with them. This analysis is detailed in Supplementary File 2 Table S6. In agreement with the remaining genome annotation, 82% of the genes encoding enzymes that catalyze these reactions were annotated based on *A. thaliana* genes' annotations. The remaining 18% were annotated by gene records of organisms not accounted for in the annotation workflow. KEGG pathways with more unique reactions are presented in Figure 2. Spontaneous reactions and reactions not associated with any pathway were not considered.

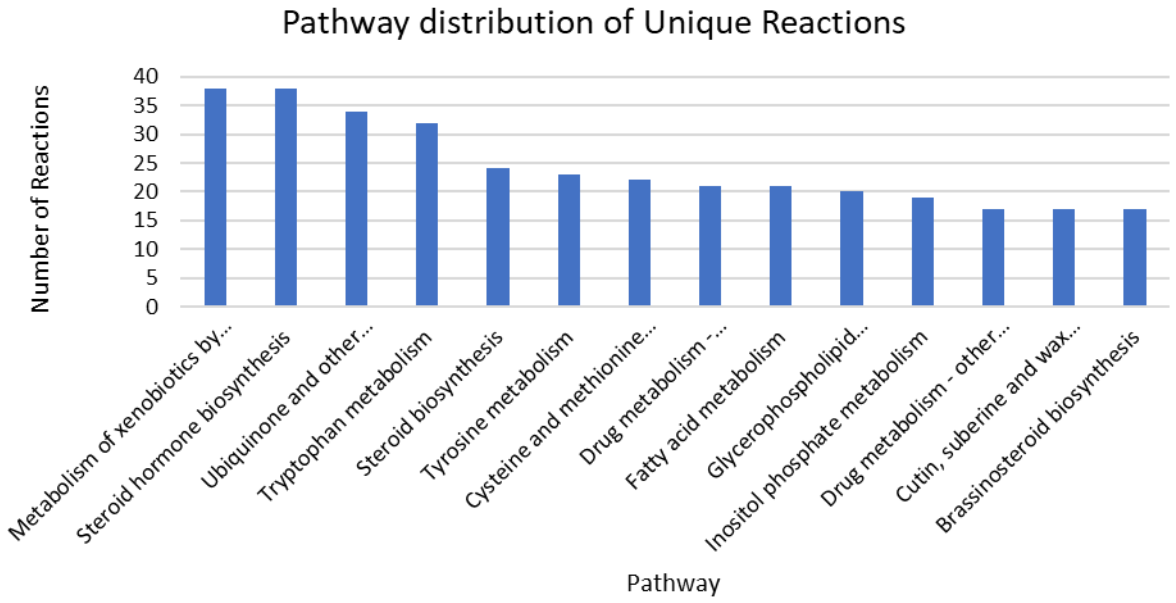

Figure 2: Pathway distribution of reactions included the cork oak model, and not in the A. thaliana and Maize models.

The unique reactions identified are associated with 122 different KEGG pathways, from which 54 had five or less unique reactions. "Steroid hormone biosynthesis", "Metabolism of xenobiotics by cytochrome P450", "Ubiquinone and other terpenoid-quinone biosynthesis", and "Tryptophan metabolism" are the pathways with more unique reactions, having 32 to 38 reactions unavailable in the maize and *A. thaliana* models.

1.3 Tissue-specific models

Transcriptomics data was integrated into the generic model using *troppo* (Ferreira et al., 2020) to obtain tissue-specific models. Leaf, Inner bark, Reproduction or traumatic Phellogen, and Virgin Phellogen were selected because of their influence in tree growth and cork production. Different biomass formulations for each tissue (Figure 3) were determined, according to experimental data available in the literature and other plant GSM models. The biomass for Reproduction and Virgin Phellogen was considered to be the same. The detailed biomass composition is available in Supplementary File 3.

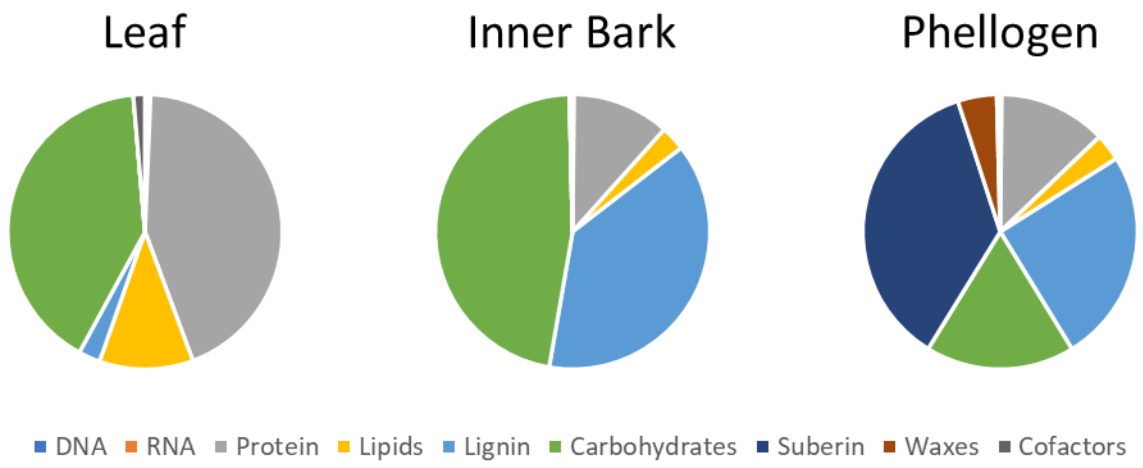

Figure 3: Biomass composition of the leaf, inner bark, and phellogen determined using data retrieved from other GSM models and available experimental data. The biomass composition and the respective data sources are detailed in Supplementary File 3.

As explained in the Materials and Methods section, the biomass was formulated by creating "e-Metabolites", representing the macromolecular composition of each tissue. Each macromolecule (e.g. e-RNA) is associated with a reaction responsible for producing it from its precursors (e.g. ATP, GTP, CTP, UTP).

The leaf macromolecular contents were determined using *A. thaliana* models (Dal'Molin et al., 2010; Arnold and Nikoloski, 2014). The cell wall sugar content was included in the e-Carbohydrate composition, while lignin was included in the e-Lignin composition. The monomer contents of DNA, RNA, and protein were determined using the biomass tool, available in *merlin*. The fatty acid, lipid, and carbohydrate compositions were determined using experimental data for *Q. suber* or closely related organisms, when species-specific data was not available (Koiwai et al., 1983; Nouairi et al., 2006; Passarinho et al., 2006). The lignin, carbohydrate, suberin, and wax contents and composition in the inner bark and phellogen were determined using available experimental data (Pereira, 1988b; Lourenço et al., 2016).

The Cofactors component includes a set of universal cofactors and vitamins (Xavier et al., 2017). These compounds were included in the biomass of each tissue. Nevertheless, the leaf "e-Cofactor" reaction also comprises a set of pigments, such as chlorophylls and carotenoids, determined according to experimental data (Garcia-Plazaola, 1997).

The formulation of tissue-specific biomass composition is critical to obtain tissue-specific models, and predict each tissue's metabolic behaviour and the interactions among them. The leaf is mainly composed of carbohydrates and protein. The inner bark also presents high amounts of carbohydrates and a significant content of lignin. Phellogen also presents considerable amounts of these macromolecules, but suberin is the one with a higher representation.

Reactions encoded by genes identified as “not expressed” were removed from each tissue-specific model. The number of genes and reactions decreased in rather different proportions, as *troppo* operates at the reaction level (See Material and Methods section 4.3). Table 2 shows the number of reactions and metabolites in the generic model and each tissue-specific model. In all tissue-specific models, the number of exchange reactions, representing the system boundaries, is 708.

Table 2: General properties of the generic and tissue-specific GSM models.

Component	Generic Model	Leaf	Inner Bark	Phellogen (reproduction)	Phellogen (virgin)
Genes	7871	7155	7153	7125	7160
Reactions	6938	5226	5258	5298	5227
• **Metabolic**	6230	4536	4563	4605	4535
• **Transport**	1269	977	940	1013	959
• **Extracellular**	176	121	124	141	140
• **Plastid**	1355	984	954	937	921
• **Cytoplasmic**	1644	1162	1189	1181	1185
• **Mitochondria**	722	537	553	523	524
• **Peroxisome**	296	223	224	205	204
• **Vacuole**	130	89	76	94	90
• **Endoplasmic Reticulum**	583	381	445	454	454
• **Golgi Apparatus**	55	44	45	42	42
Metabolites	6481	5165	5234	5240	5238

The number of metabolic reactions is quite similar among the four tissue-specific models. Nevertheless, the reproduction phellogen model presents a higher number of transport reactions than the remaining ones. The number of reactions in the leaf model's plastid is slightly higher than in the other models, while in the endoplasmic reticulum, the opposite behavior can be observed.

The number of metabolic reactions associated with each pathway is available in Supplementary File 2 Tables S8 The number of reactions associated with the “Arachidonic acid metabolism”,

“Brassinosteroid biosynthesis”, and “Drug metabolism – cytochrome P450” pathways are higher in the inner bark and phellogen than in the leaf. The “Cutin, suberin and wax biosynthesis” pathway is mostly expressed in the phellogen. As expected, the “Carotenoid biosynthesis” pathway is present in the leaf, while in the remaining tissues, the number of reactions is more limited. No significant differences were identified between the number of reactions in each pathway between the reproduction and virgin phellogen models.

1.4 Diel Multi-tissue model

A diel multi-tissue metabolic model was generated to analyze the metabolic interactions between leaf, inner bark, and phellogen at the two phases of the diel cycle: light (day) and dark (night). Common pools were created to connect the different tissues for the light and dark phases. The diel multi-tissue GSM model comprises 28,464 reactions and 29,490 metabolites. The multi-tissue GSM model allows the uptake of minerals, such as HNO_3 and H_3PO_4 through the inner bark. The light/dark uptake ratio of nitrate was constrained to 3:2, as suggested in other diel GSM models (Cheung et al., 2014; Shaw and Cheung, 2018). The exchange of oxygen and carbon dioxide is only allowed in the leaves, where photons can be uptake in the light phase. Flux variability analysis (FVA) and parsimonious flux balance analysis (pFBA) were computed, setting as objective function the minimization of photons uptake for photorespiration conditions. Figure 4 depicts both the relevant transport reactions between tissues, and the metabolites stored between the light and dark phases. The suberin synthesis pathways are also portrayed. The detailed simulation results are available in Supplementary File 4.

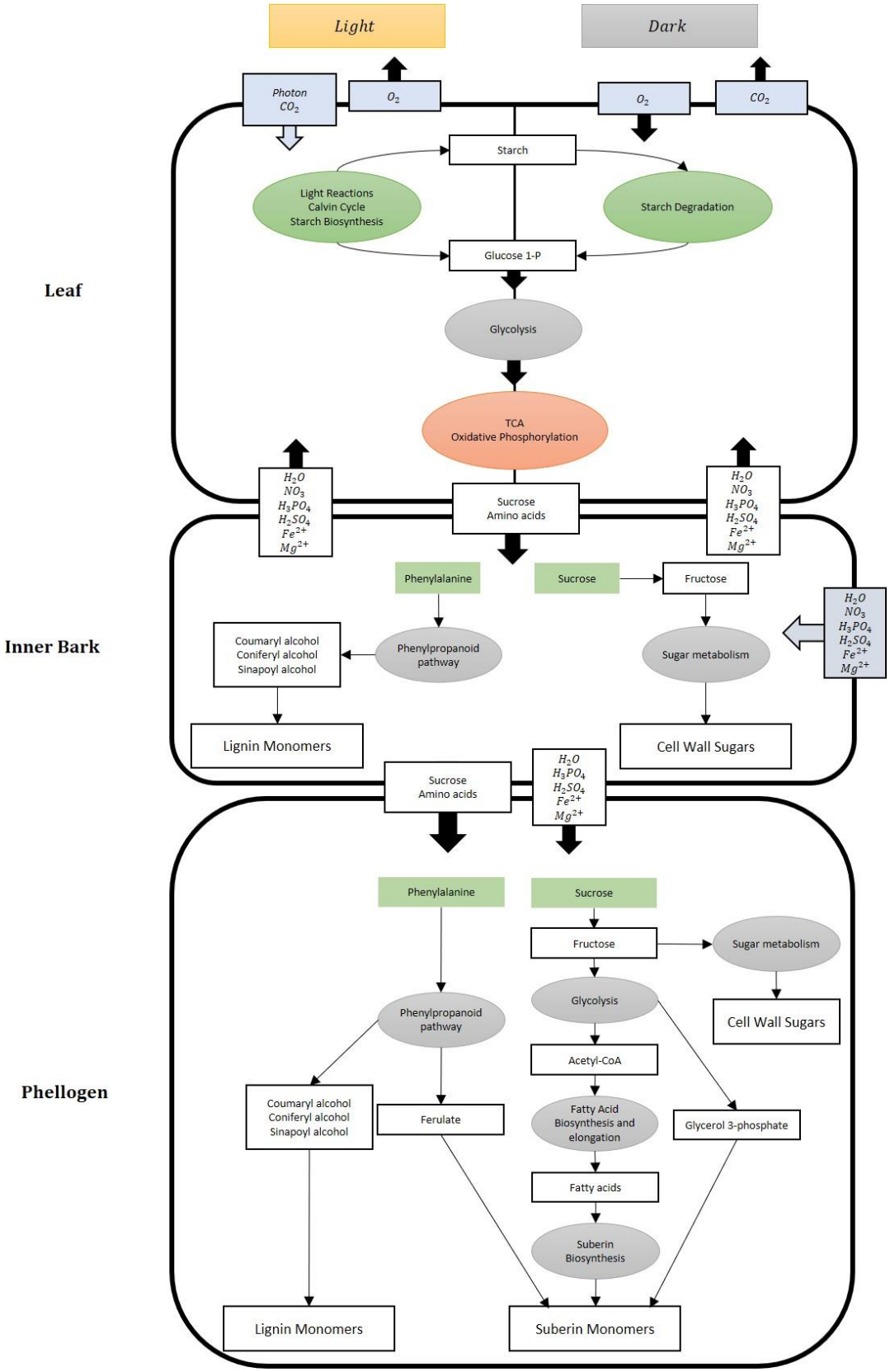

Figure 4: Schematic representation of the metabolic routes towards cork formation. Photons and gases exchanges take place in the leaf, while the uptake of inorganic ions was assumed to happen in the inner bark. Sucrose and amino acids produced in the leaf are transported into the inner bark, and then to the phellogen, where are used for the suberin and lignin biosynthesis (besides the other biomass components). In this model the inner bark metabolic processes are omitted and only the transporter and frontier role is considered. The representation is based on pFBA and FVA predictions.

2 Discussion

2.1 Comparison with other models

The number of plant GSM models has exponentially increased in the last decade. Several models have been developed and updated to obtain tissue-specific and multi-tissue GSM models. Dynamic approaches have also been applied to these models to study plant development over time.

A. thaliana and *Z. mays* are reference organisms for C3 and C4 plants; thus, several models for these organisms are available. Nevertheless, such models for less studied plants are becoming available (Grafahrend-Belau et al., 2013; Yuan et al., 2016; Pfau et al., 2018; Moreira et al., 2019; Shaw and Maurice Cheung, 2019). The iEC7871 was compared with two available models that provide KEGG identifiers, which facilitates the comparison.

The cork oak model presents 1117 reactions and 945 metabolites not available in the other two models. Interestingly, the annotation of the genes associated with these reactions was based on *A. thaliana*'s homologs. The information available in SwissProt for such species is exceptionally high, while this information is scarce for woody plants, which may lead to wrong and missing gene annotations as *Q. suber* is significantly different from *A. thaliana.* Another obstacle is the presence of incomplete EC numbers. For instance, over 300 genes were initially annotated with incomplete EC numbers associated with Cytochromes P450.

The KEGG Pathways associated with these reactions were also analyzed. The pathways with the higher number of unique reactions are associated with the metabolism of steroids, quinones, and drugs. Also, the "Tryptophan metabolism" and "Tyrosine metabolism" pathways include the secondary metabolism of these amino acids, such as the quinolinic acid production, dopamine, and its derivates. The high number of reactions and metabolites only available in the cork oak model is most likely associated with three main reasons: the approach followed to develop the GSM model; the metabolic differences between the three species; the inclusion of pathways associated with secondary plant metabolism.

2.2 Secondary Metabolism Pathways

Plants produce a wide range of compounds through their secondary metabolism, whose function includes defence against abiotic and biotic stress or beneficial interactions with other organisms. Many secondary metabolites have central roles in the pharmaceutical, cosmetics, perfume, dye, and flavour industries. Despite the recent advances in the investigation of plant secondary metabolism, detailed knowledge of these pathways is restricted to a few species (Isah, 2019), such as *A. thaliana*,

Oryza sativa and *Zea mays*. These pathways are not always complete or available in biological databases, implying an additional effort to find information regarding genes, reactions, and metabolites associated with secondary plant metabolism. Although there is still much to learn about these compounds and respective biosynthetic pathways, genome-scale modelling can provide insights into a given organism's potential to produce secondary metabolites. The cork oak model reconstructed in this work includes reactions associated with the biosynthesis and metabolism of several secondary metabolites, steroids, and drugs.

The "Metabolism of xenobiotics by cytochrome P450" includes 38 reactions not found in the *A. thaliana* and maize models considered in the Results section. This pathway represents a set of reactions, mostly associated with the cytochrome P450, responsible for the response to the presence of toxic xenobiotics. The model includes the complete routes for the degradation of benzopyrene, aflatoxin B1, and nicotine-derived xenobiotics. It was already described that the cytochrome P450 family is highly upregulated in edveloping phellem tissues (Lopes et al., 2020). The glutathione transferase enzyme (EC: 2.5.1.18), responsible for the catalysis of several reactions in this pathway due to its broad specificity, is encoded by 112 genes in the model.

The model presents all the necessary reactions to produce jasmonic acid and its derivates, usually called jasmonates, through the "alpha-Linoleic acid metabolism" pathway. These plant hormones are associated with the regulation of growth and developmental processes, stomatal opening, inhibition of Rubisco biosynthesis, and nitrogen and phosphorus uptake (Wasternack and Hause, 2013; Ruan et al., 2019). The "Steroid biosynthesis" pathway is essentially complete, enabling the production of steroids like ergosterol, cholesterol, and stigmasterol. The enzymes associated with the synthesis of calcitetrol and secalciferol (animal hormones) are the only ones whose encoding gene were not found in the genome of *Q. suber*. Although these hormones are usually synthesized by animals, they were already identified in a few plants (Aburjai et al., 1998; Boland et al., 2003; Jäpelt and Jakobsen, 2013).The biosynthetic pathway for the synthesis of gibberellins is also available in the model. Although the regulatory effect of hormones and steroids cannot be quantitively evaluated in stoichiometric models, the presence of their biosynthetic pathways, connected to the core network, can be used as a source of information regarding the potential to produce determined hormone or steroid.

Although this model contains a considerable amount of information regarding secondary metabolism, further curation of these pathways would improve the connectivity of the model and reduce the existing gaps.

2.3 Tissue-specific models

A generic metabolic model comprises all the reactions catalyzed by all the enzymes encoding in an organism's genome. However, this approach does not account for the regulatory network present in each organ, tissue, or cell in different environmental conditions. The integration of transcriptomics data in a GSM model allows obtaining metabolic models closer to the *in vivo* phenotype of the respective tissue or condition.

Transcriptomics data for the leaf, inner bark, and phellogen were integrated into the generic model, originating three tissue-specific models. The biomass of each tissue was determined using data available in literature and other GSM models. The leaf macromolecular composition was mostly based on previously published GSM models. Quantitative information regarding the leaf composition of *Q. suber* would be useful to increase the model's reliability and improve predictions. The composition of the inner bark and the phellogen was based on experimental data available for *Q. suber*. While the inner bark is mostly composed of lignin and carbohydrates, the phellogen also comprises suberin and waxes. Suberin is composed of glycerol, ferulic acid, and diverse alkanols, fatty acids, hydroxy acids, ω-hydroxy acids and α,ω-dicarboxylic acids. Waxes are composed by terpenes, such as friedelin, and sterols, like β-sitosterol. Although the cork bark contains tannins (mostly ellagitannins) (Cadahía et al., 1998), these compounds were not included in the biomass since their biosynthetic pathways are poorly understood and not available in the used biological databases.

The number of reactions and metabolites on each subcellular compartment was compared across the three tissues. The number of transport reactions in the Reproduction Phellogen is visibly higher than in the other tissues, including the Virgin Phellogen. These reactions are essentially associated with the transport of steroids and hormones. Based on BLAST searches against TCDB, the genes associated with these reactions are auxin (phytohormone required for cell division (Perrot-Rechenmann and Napier, 2005) efflux pumps. An analysis of the pathways available in the Reproduction and Virgin Phellogen models did not find significant differences between them. The cork oak phellogen's transcriptional profile allows inferring that the flavonoid route is favoured in bad quality cork, while the lignin and suberin production pathways are preferred in good quality cork production (Teixeira et al., 2018). A recent study reported a higher expression of genes associated with fatty acid biosynthesis and elongation in reproduction cork (Lopes et al., 2020). Hence, a different approach should be followed to identify these differences in the metabolic models. First, a biomass composition specific to the virgin phellogen should be defined, including the tannin content. Although it was reported that the macromolecular composition of reproduction and virgin cork could be similar, the precursors of suberin and extractives are significantly different (Pereira, 1988a). The comparison between the two models should be based on flux prediction in specific pathways, not on the number of reactions.

2.4 Suberin Biosynthesis

The synthesis of the aliphatic suberin monomers is associated with the KEGG's "Cutin, suberin, and wax biosynthesis" pathway. Fatty acids are exported from the plastid and imported into the endoplasmic reticulum, where these can be elongated to origin long-chain and very-long-chain fatty acids. Cytochromes P450, especially of the CYP86A family, are responsible for catalyzing several reactions that synthesise suberin monomers. CYP86A1 and CYP704B1 are responsible for the ω-hydroxylation of LCFA, while CYP86B1 acts on VLCFA.

Epoxide hydrolases (EC: 3.3.2.10), ω-oxo-fatty acid dehydrogenases (EC: 1.2.1.-), ω-hydroxy-fatty acid dehydrogenases (EC: 1.1.1.-), and peroxygenases (EC: 1.11.2.3) can act in the ω-hydroxy acid in different combinations, to originate the diverse α,ω-dicarboxylic, poly α,ω-dicarboxylic, and polyhydroxy-fatty acids. The genome annotation and the model's manual curation allowed identifying 38 genes encoding epoxide hydrolase, five encoding ω-hydroxy-fatty acid dehydrogenase, and six encoding peroxygenase. Nevertheless, no genes encoding ω-oxo-fatty acid dehydrogenase were available in any database; thus, the reactions catalyzed by this enzyme were added without gene association.

The polymerization process is still not completely understood but is likely involves esterification reactions through glycerol-3-phosphate acyltransferases, producing acylglycerol esters. These are secreted through the Golgi secretory pathway and ABC transporters and incorporated in the suberin glycerol polyester (Vishwanath et al., 2015). The aliphatic suberin polyester assembly was simplified in the model and represented through a reaction converting the determined precursors (Supplementary File 3) into the macromolecular representation of suberin: "e-Suberin".

As mentioned before, the bark also contains a considerable amount of lignin (aromatic suberin). The synthesis of the respective monomers (guaciacyl lignin, hydroxyphenyl lignin, and syringyl lignin) is described in the "Phenylpropanoid biosynthesis" pathway. Phenylalanine is the main precursor of this pathway that occurs mostly in the cytoplasm. It is converted to cinnamic acid by phenylalanine ammonia-ligase (EC: 4.3.1.24). Cinnamate monoxigenase, caffeate methyltransferase, and ferulate-5-hydroxylase are responsible for the successive conversion of cinnamic acid to coumarate, ferulate, hydroxyferulate, and sinapic acid. These metabolites can be converted into lignin monomers in a four-step process catalyzed by coumarate-CoA ligase (EC: 6.2.1.12), cinnamoyl-CoA reductase (EC: 1.2.1.44), cinammyl-alcohol dehydrogenase (EC: 1.1.1.95), and finally peroxidase (EC: 1.11.1.7). The former enzyme is responsible for the assembly of the monomers into an aromatic polymer in the cell

wall. The linkages between aliphatic and aromatic suberin was not included in the model since these macromolecules are represented in the biomass separately.

2.5 Diel multi-tissue model

Tissues and organs of multicellular organisms do not perform their metabolic functions individually. Instead, they interact with each other by exchanging sugars, amino acids, hormones, and others. In plants, the simplest example is the sucrose formation in the leaves and its transport across all other organism tissues. Multi-tissue GSM models allow analysing the dependencies between tissues in terms of biomass precursors, carbon skeletons, nitrogen, sulfur, phosphorus sources, and the energy required for their translocation and proton balance (Bordbar et al., 2011; Zakhartsev et al., 2016). The introduction of light and dark phases allows predicting features that are not possible in continuous light models, such as the accumulation and utilization of carboxylic acids (Cheung et al., 2014).

The leaf, inner bark, and phellogen models were merged into a multi-tissue GSM model. Light and dark phases were considered to present the metabolic differences that occur in these two phases. Other tissues or organs, such as the root and xylem, could also be considered and included in the multi-tissue model. Nevertheless, this work's scope is related to the metabolism of cork precursors, and including more tissues would increase the complexity of the model unnecessarily.

The most significant differences between the light and dark phases of metabolism were identified in the leaf. More than half of the storage reactions with predicted flux are localized in this tissue (Supplementary File 4) due to the photosynthesis and photorespiration pathways.

In the light phase, carbon dioxide is fixed through the Calvin Cycle, forming carbohydrates, and consequently, the carbon skeletons for the synthesis of the remaining biomass components. A fraction of the carbohydrates produced here (mainly starch) is stored and used as energy source when needed. The storage of starch is more efficient in the model than using sucrose, glucose, or fructose. Whereas starch is mobilized in plastids, soluble sugars are accumulated in the vacuole, implying an energetic cost to its transport. The model produced malate in the dark phase, which is used at day by the malic enzyme to produce pyruvate and obtain reductive power (NADPH). Citrate was also produced at night through the tricarboxylic cycle (TCA) and used to feed the diurnal TCA and provide carbon skeletons for the amino acid metabolism. This behaviour is in agreement with the previously published diel GSM model for *A. thaliana* (Cheung et al., 2014).

Nitrate provided by the inner bark is transported into the leaf and is then converted into ammonia. In the plastids, ammonia is used to incorporate nitrogen through the glutamine synthetase-glutamate synthase pathway.

Sucrose and amino acids are exported from the leaf and then imported by proton symport by the inner bark. The proton balance of the common pool is maintained by plasma membrane ATPases. In the

inner bark, sucrose is used as an energy and carbon source to produce the biomass components. As mentioned before, the inner bark biomass is mainly constituted by carbohydrates, lignin, and protein. Hence, pathways associated with the synthesis of cell wall precursors ("Amino sugar and amino nucleotide metabolism", "Starch and sucrose metabolism", and "Phenylpropanoid biosynthesis") are the most relevant in this tissue.

The remaining sucrose and amino acids from the leaves are transported to the phellogen. The reversible sucrose synthase activity allows sucrose conversion into UDP-glucose, which is used to produce other cell wall components, and fructose that follows the glycolytic pathway to produce pyruvate. A fraction of it is transported into the plastid and converted by pyruvate dehydrogenase into acetyl-CoA, the fatty acids precursor. As mentioned above, fatty acids are elongated and suffer a series of hydroxylation, epoxidation, and peroxidation reactions in the endoplasmic reticulum. Ferulate, another precursor of suberin, is produced through the "Phenylpropanoid biosynthesis" pathway, with phenylalanine serving as a precursor. The glycerol 3-phosphate used for the esterification reactions is produced from glycerone phosphate produced through glycolysis.

The diel multi-tissue GSM model developed in this work is a useful framework capable of providing insights into the metabolism of *Q. suber* at a global level. It can be used to study the biosynthetic pathways of suberin, lignin, waxes, and many other compounds.

In recent years, plant multi-tissue GSM models have been used for dynamic approaches, such as dynamic flux balance analysis (dFBA). An *A. thaliana* model was used to study the impact of carbon and nitrogen metabolism in the biomass of leaf and root (Shaw and Cheung, 2018). The dFBA approach was also applied to the barley model to study source and sink organs' metabolic behaviour through the generative phase of the organism (Grafahrend-Belau et al., 2013). This approach can be applied to the models developed in this work to study, for example, the metabolic evolution of the phellogen regeneration after harvest.

3 Conclusion

In this work, we present the first genome-scale metabolic model of a woody plant. The iEC7871 was based on knowledge retrieved from genomic information, biological databases, and literature. It can simulate the Cork Oak tree's metabolic behaviour in phototrophic and heterotrophic conditions, as well as photorespiration. This model comprises the pathways of the central metabolism and several pathways associated with the secondary metabolism reproducying the formation of the major constituints of cork.

The integration of transcriptomics data allowed obtaining tissue-specific models for the leaf, inner bark, and phellogen. These models were merged to obtain a multi-tissue GSM model that comprises the diel cycle's dark and light phases.

The metabolic models developed in this work can be used as a tool to analyze and predict the metabolic behaviour of the tree and evaluate its metabolic potential. Metabolic modelling methods can be applied, including dynamic approaches, to study the changes in this tree's metabolism over time and environment.

4 Materials and Methods

4.1 Software

merlin v4 was used to support the reconstruction process. This user-friendly tool presents a graphical interface and a set of plug-ins that assist and accelerate the reconstruction of GSM models, namely in the genome annotation, network assembly, and manual curation.

The COBRApy v0.20.0 (Ebrahim et al., 2013) tool was used to perform all simulations and analyses of the GSM model, as well as generate the diel multi-tissue model. The simulations were performed using the CPLEX v128.0.0 solver.

The *Troppo* (Ferreira et al., 2020) python tool was used to integrate the transcriptomics data in the metabolic model, originating tissue-specific models.

The FastQC (Andrews, 2015), Sickle, Bowtie2, FeatureCounts, and edgeR (Robinson et al., 2009) tools and packages were used to process the transcriptomics data to be integrated into the GSM model.

4.2 Metabolic Reconstruction

The "Automatic workflow" tool, available in *merlin*, was used to perform the genome annotation by assigning EC numbers to enzyme encoding genes based on BLAST searches against SwissProt and TrEMBL. The details of the genome annotation are available in Supplementary Material S1.

A draft metabolic network was assembled by loading KEGG's metabolic information and integrating the genome annotation. KEGG reactions associated with enzymes identified in the genome annotation stage were included in the model, as well as spontaneous reactions.

Transport reactions were automatically generated using the TranSyT tool. Nevertheless, additional transport reactions were added if it is reported in the literature or if necessary, for the model functionality.

The subcellular location of proteins was predicted using the WolfPsort tool. Additionally, the LocTree and ChloroP 1.1 tools were used to verify protein locations through the manual curation stage. The mitochondrial and thylakoidal intermembrane regions were added manually in the photosynthesis and oxidative phosphorylation reactions.

The leaf, inner bark, and phellogen biomass compositions were based on previously published plant GSM models and available literature. The biomass compositions were organized in macromolecules or cell structures, labeled as "e-Metabolites". Suberin was only accounted for in inner bark and phellogen. The "e-Cofactor" content includes cofactors, vitamins, and pigments (only included in leaf biomass). The energy requirements were retrieved from the x model. The complete biomass composition is available in Supplementary Material S2.

Through the whole process, literature, and biological databases, namely NCBI, KEGG, MetaCyc, and BRENDA, were consulted to retrieve information regarding metabolic, genomic, and enzymatic information.

The validation of the GSM model started by assuring the biomass formation, which was performed using the BioISO tool available in *merlin*.

The following validation approaches were applied to the model:

1. Growth without photons, carbon, nitrogen, phosphorus, and sulfur sources.
2. Futile cycles and stoichiometrically balanced cycles.
3. Growth rate assessment.
4. Growth with different elemental sources.
5. Capacity to present flux through photosynthesis, respiration, and photorespiration.

The model validation was performed using FBA, pFBA, and FVA, available in COBRApy. All the metabolic models developed in this work are available at https://cutt.ly/qsubermodels.

4.3 Tissue-Specific and multi-tissue Models

Transcriptomics data from different tissues was used to obtain tissue-specific GSM models. The quality of the transcriptomics data was analyzed using the FastQC software. After trimming the data with Sickle, Bowtie2 was used to align the trimmed FASTQ files against the reference genome. FeatureCounts was used to count the mapped reads to each gene. After filtering and normalizing the data with the edgeR library, datasets with the normalized counts of each gene were obtained. These files were used as input for *troppo*, together with the generic GSM model, to generate tissue-specific models. Troppo was ran using the CPLEX solver and the "Fastcore" algorithm. A *troppo*'s integration strategy was developed to guarantee that the tissue-specific models are capable of producing biomass. The median of each dataset was used as threshold, and the remaining parameters were kept as default. *Troppo* identifies the reactions that should be removed from the model, based on the expression of the genes included in the GPR association of each reaction. Nevertheless, if a reaction is kept in the model, the respective genes are not removed. The carotenoid biosynthesis has two branches, one associated with trans-Phytofluene and the other with 15,9-dicis-Phytofluene. The enzymes present in the two branches are similar (phytoene desaturase and zeta-carotene desaturase). However, one of them is dependent on the photosynthesis, since it unbalances the plastoquinone/plastoquinol ratio. The other one is stoichiometrically balanced since it involves successive oxidation and reduction of these quinones. Hence, the presence of the second branch was guaranteed in the leaf model, since it is essential for the carotenoid production through the dark phase.

The tissue-specific models obtained using *troppo* were adapted to account for the light and dark conditions, by duplicating all reactions and metabolites. Reactions converting sugars (starch, sucrose, glucose, fructose, malate, citrate, quercitol, and quinate), 18 amino acids, and nitrate between light and dark were added to the model, based on a previously published approach (Cheung et al., 2014). To build the multi-tissue model, the diel tissue-specific models were merged, and common pools between leaf and inner bark (common pool 1), and inner bark and phellogen (common pool 2) were created. Since the root was not accounted for, the uptake of water, nitrogen, sulfur, phosphorus, and magnesium sources were included in the inner bark.

4.4 Simulations

The FBA (Flux Balance Analysis) computational approach (Orth et al., 2010) has been extensively used as a mathematical framework for studying several metabolic networks (Duarte et al., 2007; Orth et al., 2011; McCloskey et al., 2013). Based on linear programming, the FBA approach uses an under-determined system of linear equations to calculate the flow of metabolites through the metabolic network, for a given objective. The optimization is subject to a set of constraints, such as the reversibility and bounds of reactions, and stoichiometry of metabolites. Besides the constraints, FBA also requires the definition of a phenotype as the objective function that depends on the scope of the work. Usually, the objective function is the maximization of biomass production or a by-product, or the minimization of substrate uptake.

Two different approaches were used to perform simulations with pFBA. The first was to fix the biomass formation to 0.11 h^{-1}, and set the minimization of the photon uptake as the objective function. In the second approach, the photon uptake was fixed to 100 $mmol_{photon}.gDW^{-1}.h^{-1}$, maximizing the biomass reaction. Photorespiration was simulated by constraining the carboxylation/oxygenation flux ratio of Rubisco (reactions R00024 and R03140) to 3:1. The simulations with the diel multi-tissue GSM model were performed using the restrictions applied in photorespiration, with the additional constraint of the nitrate light/dark uptake ratio, which was settled to 3:2.

4.5 Accession Numbers

The genome sequence of *Quercus suber* was retrieved from the NCBI database with the assembly accession number GCF_002906115.1. The transcriptomics data was retrieved from the EBI database(Li et al., 2015) (Accession PRJNA392919 for Leaf and Inner Bark; Accession PRJEB33874 for Phellogen).

References

Aburjai T, Al-Khalil S, Abuirjeie M (1998) Vitamin D3 and its metabolites in tomato, potato, egg plant and zucchini leaves. Phytochemistry. doi: 10.1016/S0031-9422(98)00246-5

Acácio V, Holmgren M, Jansen PA, Schrotter O (2007) Multiple recruitment limitation causes arrested succession in Mediterranean cork oak systems. Ecosystems. doi: 10.1007/s10021-007-9089-9

Agren R, Mardinoglu A, Asplund A, Kampf C, Uhlen M, Nielsen J (2014) Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. Mol Syst Biol. doi: 10.1002/msb.145122

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215**: 403–410

Andrews S (2015) FASTQC A Quality Control tool for High Throughput Sequence Data. Babraham Inst.

Apweiler R, Martin MJ, O'Donovan C, Magrane M, Alam-Faruque Y, Antunes R, Barrell D, Bely B, Bingley M, Binns D, et al (2011) Ongoing and future developments at the Universal Protein Resource. Nucleic Acids Res **39**: D214–D219

Arnold A, Nikoloski Z (2014) Bottom-up metabolic reconstruction of arabidopsis and its application to determining the metabolic costs of enzyme production. Plant Physiol. doi: 10.1104/pp.114.235358

Boland R, Skliar M, Curino A, Milanesi L (2003) Vitamin D compounds in plants. Plant Sci. doi: 10.1016/S0168-9452(02)00420-X

Bordbar A, Feist AM, Usaite-Black R, Woodcock J, Palsson BO, Famili I (2011) A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. BMC Syst Biol. doi: 10.1186/1752-0509-5-180

Cadahía E, Conde E, Fernández De Simón B, García-Vallejo MC (1998) Changes in Tannic Composition of Reproduction Cork Quercus suber throughout Industrial Processing. J Agric Food Chem. doi: 10.1021/jf9709360

Cheung CYM, Poolman MG, Fell DA, Ratcliffe RG, Sweetlove LJ (2014) A Diel Flux Balance Model Captures Interactions between Light and Dark Metabolism during Day-Night Cycles in C 3 and Crassulacean Acid Metabolism Leaves. Plant Physiol **165**: 917–929

Cheung CYM, Williams TCR, Poolman MG, Fell DA, Ratcliffe RG, Sweetlove LJ (2013) A method for accounting for maintenance costs in flux balance analysis improves the prediction of plant cell metabolic phenotypes under stress conditions. Plant J **75**: 1050–61

Chung BKS, Lakshmanan M, Klement M, Mohanty B, Lee DY (2013) Genome-scale in silico modeling and analysis for designing synthetic terpenoid-producing microbial cell factories. Chem Eng Sci. doi: 10.1016/j.ces.2012.09.006

Dal'Molin CG de O, Quek LE, Palfreyman RW, Brumbley SM, Nielsen LK (2010) AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. Plant Physiol. doi: 10.1104/pp.109.148817

Dias O, Rocha I (2015) Systems Biology in Fungi. *In* R Paterson, ed, Mol. Biol. Food Water Borne Mycotoxigenic Mycotic Fungi. CRC Press, Boca Raton, pp 69–92

Dias O, Rocha M, Ferreira EC, Rocha I (2015) Reconstructing genome-scale metabolic models with merlin. Nucleic Acids Res **43**: 3899–3910

Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, Srivas R, Palsson BØ (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc Natl Acad Sci U S A **104**: 1777–82

Ebrahim A, Lerman JA, Palsson BO, Hyduke DR (2013) COBRApy: COnstraints-Based Reconstruction and Analysis for Python. BMC Syst Biol. doi: 10.1186/1752-0509-7-74

Ferreira J, Vieira V, Gomes J, Correia S, Rocha M (2020) Troppo - A Python Framework for the Reconstruction of Context-Specific Metabolic Models. Adv Intell Syst Comput. doi: 10.1007/978-3-030-23873-5_18

Fong SS, Burgard AP, Herring CD, Knight EM, Blattner FR, Maranas CD, Palsson BO (2005) In silico design and adaptive evolution of Escherichia coli for production of lactic acid. Biotechnol Bioeng. doi: 10.1002/bit.20542

Garcia-Plazaola J (1997) Seasonal changes in xanthophyll composition and photosynthesis of cork oak (Quercus suber L.) leaves under Mediterranean climate. J Exp Bot. doi: 10.1093/jexbot/48.314.1667

Graça J, Pereira H (2004) The periderm development in Quercus suber. IAWA J. doi: 10.1163/22941932-90000369

Grafahrend-Belau E, Junker A, Eschenröder A, Müller J, Schreiber F, Junker BH (2013) Multiscale

metabolic modeling: Dynamic flux balance analysis on a whole-plant scale. Plant Physiol. doi: 10.1104/pp.113.224006

Jäpelt RB, Jakobsen J (2013) Vitamin D in plants: A review of occurrence, analysis, and biosynthesis. Front Plant Sci. doi: 10.3389/fpls.2013.00136

Jerby-Arnon L, Pfetzer N, Waldman YY, McGarry L, James D, Shanks E, Seashore-Ludlow B, Weinstock A, Geiger T, Clemons PA, et al (2014) Predicting cancer-specific vulnerability via data-driven detection of synthetic lethality. Cell. doi: 10.1016/j.cell.2014.07.027

Kim HN, Jin HY, Kwak MJ, Khaine I, You HN, Lee TY, Ahn TH, Woo SY (2017) Why does Quercus suber species decline in Mediterranean areas? J Asia-Pacific Biodivers. doi: 10.1016/j.japb.2017.05.004

Koiwai A, Suzuki F, Matsuzaki T, Kawashima N (1983) The fatty acid composition of seeds and leaves of Nicotiana species. Phytochemistry **22**: 1409–1412

Lakshmanan M, Zhang Z, Mohanty B, Kwon J-Y, Choi H-Y, Nam H-J, Kim D-I, Lee D-Y (2013) Elucidating rice cell metabolism under flooding and drought stresses using flux-based modeling and analysis. Plant Physiol **162**: 2140–50

Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N, Lopez R (2015) The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Res. doi: 10.1093/nar/gkv279

Lopes ST, Sobral D, Costa B, Perdiguero P, Chaves I, Costa A, Miguel CM (2020) Phellem versus xylem: Genome-wide transcriptomic analysis reveals novel regulators of cork formation in cork oak. Tree Physiol. doi: 10.1093/treephys/tpz118

Lourenço A, Rencoret J, Chemetova C, Gominho J, Gutiérrez A, Del Río JC, Pereira H (2016) Lignin composition and structure differs between Xylem, Phloem and phellem in quercus suber L. Front Plant Sci. doi: 10.3389/fpls.2016.01612

McCloskey D, Palsson BØ, Feist AM (2013) Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli. Mol Syst Biol **9**: 661

Moreira TB, Shaw R, Luo X, Ganguly O, Kim HS, Coelho LGF, Cheung CYM, Williams TCR (2019) A genome-scale metabolic model of soybean (Glycine max) highlights metabolic fluxes in seedlings. Plant Physiol. doi: 10.1104/pp.19.00122

Nouairi I, Ammar W Ben, Youssef N Ben, Daoud DBM, Ghorbal MH, Zarrouk M (2006) Comparative

study of cadmium effects on membrane lipid composition of Brassica juncea and Brassica napus leaves. Plant Sci **170**: 511–519

Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, Palsson BØ (2011) A comprehensive genome-scale reconstruction of Escherichia coli metabolism--2011. Mol Syst Biol **7**: 535

Orth JD, Thiele I, Palsson BO (2010) What is flux balance analysis? Nat Biotechnol **28**: 245–248

Passarinho JAP, Lamosa P, Baeta JP, Santos H, Ricardo CPP (2006) Annual changes in the concentration of minerals and organic compounds of Quercus suber leaves. Physiol Plant **127**: 100–110

Pereira-Leal JB, Abreu IA, Alabaça CS, Almeida M, Almeida P, Almeida T, Amorim M, Araújo S, Azevedo H, Badia A, et al (2014) A comprehensive assessment of the transcriptome of cork oak (Quercus suber) through EST sequencing. BMC Genomics **15**: 371

Pereira H (1988a) Chemical composition and variability of cork from Quercus suber L. Wood Sci Technol **22**: 211–218

Pereira H (1988b) Chemical composition and variability of cork from Quercus suber L. Wood Sci Technol **22**: 211–218

Pereira H (2015) The Rationale behind Cork Properties: A Review of Structure and Chemistry. BioResources. doi: 10.15376/biores.10.3.Pereira

Perrot-Rechenmann C, Napier RM (2005) Auxins. Vitam. Horm. pp 203–233

Pfau T, Christian N, Masakapalli SK, Sweetlove LJ, Poolman MG, Ebenhöh O (2018) The intertwined metabolism during symbiotic nitrogen fixation elucidated by metabolic modelling. Sci Rep. doi: 10.1038/s41598-018-30884-x

Poolman MG, Kundu S, Shaw R, Fell DA (2013) Responses to light intensity in a genome-scale model of rice metabolism. Plant Physiol. doi: 10.1104/pp.113.216762

Ramos AM, Usié A, Barbosa P, Barros PM, Capote T, Chaves I, Simões F, Abreu I, Carrasquinho I, Faro C, et al (2018) The draft genome sequence of cork oak. Sci data **5**: 180069

Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. doi: 10.1093/bioinformatics/btp616

Ruan J, Zhou Y, Zhou M, Yan J, Khurshid M, Weng W, Cheng J, Zhang K (2019) Jasmonic acid

signaling pathway in plants. Int J Mol Sci. doi: 10.3390/ijms20102479

Saha R, Suthers PF, Maranas CD (2011) Zea mays irs1563: A comprehensive genome-scale metabolic reconstruction of maize metabolism. PLoS One. doi: 10.1371/journal.pone.0021784

Seaver SMD, Bradbury LMT, Frelin O, Zarecki R, Ruppin E, Hanson AD, Henry CS (2015) Improved evidence-based genome-scale metabolic models for maize leaf, embryo, and endosperm. Front Plant Sci. doi: 10.3389/fpls.2015.00142

Shaw R, Cheung CYM (2018) A Dynamic Multi-Tissue Flux Balance Model Captures Carbon and Nitrogen Metabolism and Optimal Resource Partitioning During Arabidopsis Growth. Front Plant Sci. doi: 10.3389/fpls.2018.00884

Shaw R, Maurice Cheung CY (2019) A mass and charge balanced metabolic model of Setaria viridis revealed mechanisms of proton balancing in C4 plants. BMC Bioinformatics. doi: 10.1186/s12859-019-2941-z

Teixeira RT, Fortes AM, Bai H, Pinheiro C, Pereira H (2018) Transcriptional profiling of cork oak phellogenic cells isolated by laser microdissection. Planta. doi: 10.1007/s00425-017-2786-5

Vaz M, Maroco J, Ribeiro N, Gazarini LC, Pereira JS, Chaves MM (2011) Leaf-level responses to light in two co-occurring Quercus (Quercus ilex and Quercus suber): leaf structure, chemical composition and photosynthesis. Agrofor Syst **82**: 173–181

Vishwanath SJ, Delude C, Domergue F, Rowland O (2015) Suberin: biosynthesis, regulation, and polymer assembly of a protective extracellular barrier. Plant Cell Rep. doi: 10.1007/s00299-014-1727-z

Wasternack C, Hause B (2013) Jasmonates: Biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany. Ann Bot. doi: 10.1093/aob/mct067

Xavier JC, Patil KR, Rocha I (2017) Integration of Biomass Formulations of Genome-Scale Metabolic Models with Experimental Data Reveals Universally Essential Cofactors in Prokaryotes. Metab Eng. doi: 10.1016/j.ymben.2016.12.002

Yuan H, Cheung CYM, Poolman MG, Hilbers PAJ, Van Riel NAW (2016) A genome-scale metabolic network reconstruction of tomato (Solanum lycopersicum L.) and its application to photorespiratory metabolism. Plant J. doi: 10.1111/tpj.13075

Zakhartsev M, Medvedeva I, Orlov Y, Akberdin I, Krebs O, Schulze WX (2016) Metabolic model of

central carbon and energy metabolisms of growing Arabidopsis thaliana in relation to sucrose translocation. BMC Plant Biol. doi: 10.1186/s12870-016-0868-3

Zhang C, Hua Q (2016) Applications of genome-scale metabolic models in biotechnology and systems medicine. Front Physiol. doi: 10.3389/fphys.2015.00413